DM2 : Lemme d'Arden

Ce devoir est facultatif mais il est vivement conseillé de réfléchir aux questions 1 à 7 car le lemme d'Arden est un résultat extrêmement classique. Il est à rendre pour le 12/10. L'objectif de ce sujet est d'étudier des (systèmes) d'équations dont les inconnues sont des langages et d'en explorer quelques applications.

Partie 1 *Lemme d'Arden*

Dans cette partie, Σ est un alphabet et $K, L \in \mathcal{P}(\Sigma^\star)$ sont deux langages sur cet alphabet. On y étudie l'équation $(E) : X = KX + L$ dont l'inconnue est le langage X.

1. Montrer que le langage $K^\star L$ est solution de l'équation (E).

2. Montrer que toute solution X de (E) vérifie $K^\star L \subset X$.

Ces deux questions montrent que $K^\star L$ est la plus petite solution de (E) au sens de l'inclusion.

3. On suppose ici que $\varepsilon \notin K$. Montrer sous cette condition que toute solution X de (E) vérifie $X \subset K^\star L$. *Indication : On pourra procéder par l'absurde et considérer un mot de longueur minimale de $X \setminus K^\star L$.*

4. Déduire des questions précédentes le lemme d'Arden : Pour tout alphabet Σ, pour tous $L, K \in \mathcal{P}(\Sigma^\star)$ tels que $\varepsilon \notin K$, l'équation $X = KX + L$ admet une unique solution, à savoir $K^\star L$.

5. A-t-on toujours unicité de la solution à l'équation (E) si $\varepsilon \in K$? Justifier.

Partie 2 *Systèmes d'équations aux langages*

Dans cette partie, $\Sigma = \{a, b\}$. On note L_1 le langage des mots ayant un nombre pair de b et L_2 le langage des mots ayant un nombre impair de b.

6. Expliquer pourquoi (L_1, L_2) est solution du système d'équations suivant :

$$(S) : \begin{cases} L_1 = aL_1 + bL_2 + \varepsilon & (1) \\ L_2 = aL_2 + bL_1 & (2) \end{cases}$$

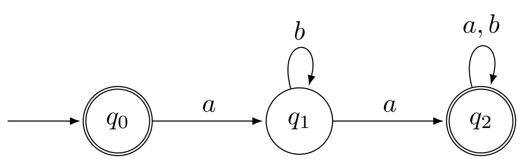
7. En utilisant le lemme d'Arden, résoudre le système (S) et en déduire une expression rationnelle pour L_1 et L_2. *Indication : utiliser le lemme d'Arden sur* (2) *puis substituer dans* (1).

8. Montrer que les expressions rationnelles $a^\star b(a + ba^\star b)^\star$ et $(a + ba^\star b)^\star ba^\star$ sont équivalentes et dénotent toutes les deux le langage L_2. *Indication : résoudre le système* (S) *d'une manière différente.*

Les techniques utilisées dans les questions précédentes offrent une méthode pour montrer que deux expressions rationnelles sont équivalentes : on montre qu'elles sont solutions d'un même système d'équations aux langages et on utilise à bon escient l'unicité offerte par le lemme d'Arden.

Partie 3 *Langage associé à un automate*

L'objectif de cette partie est de présenter une méthode permettant de déterminer le langage reconnu par un automate et reposant sur la résolution d'équations aux langages. Dans la question 9, on considère l'automate $A = (\{a, b\}, Q, q_0, F, \delta)$ suivant :



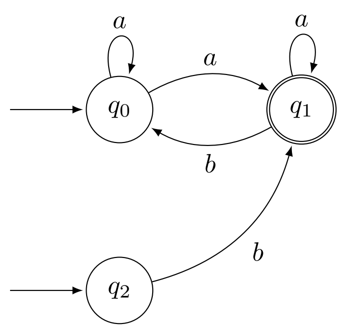
Pour $i \in \{0, 1, 2\}$, on note $L_i = \{m \in \Sigma^\star \mid \delta^\star(q_i, m) \in F\}$, qui n'est rien d'autre que le langage des mots qui font aboutir à un état final à partir de l'état q_i. Déterminer le langage reconnu par A revient ici à déterminer L_0.

9. La lecture de l'automate donne des liens entre les langages L_0, L_1 et L_2. Par exemple, $L_0 = \varepsilon + aL_1$.

   a) Déterminer un système de trois équations liant les langages L_0, L_1 et L_2.

   b) Résoudre ce système et en déduire le langage reconnu par A.

10. En utilisant une méthode similaire, déterminer le langage reconnu par l'automate suivant :



Partie 4 *Les langages reconnus sont rationnels*

L'objectif de cette partie est de donner une nouvelle preuve pour l'une des implications du théorème de Kleene : tout langage reconnaissable par un automate est rationnel. Pour ce faire, on s'appuie sur la technique de la partie 3 et sur le lemme d'Arden.

11. Montrer par récurrence sur $n \in \mathbb{N}^\star$ le théorème suivant :

    Soit $(K_{i,j})_{i,j \in [\![0,n-1]\!]}$ un n^2-uplets de langages sur un alphabet Σ dont aucun ne contient ε. Soit $(L_0, ..., L_{n-1})$ un n-uplet de langages quelconques. Alors le système d'équations :

$$X_i = \left(\sum_{j=0}^{n-1} K_{i,j} X_j \right) + L_i \text{ pour } i \in [\![0, n-1]\!]$$

    d'inconnues les langages $(X_0, ..., X_{n-1})$ admet une unique solution. De plus si les langages $K_{i,j}$ et L_i sont tous rationnels alors les composantes X_i de cette unique solution sont également des langages rationnels.

12. Soit L un langage reconnaissable par un automate. En s'inspirant de la partie 3, montrer que L est l'une des composantes d'une solution d'un système d'équations aux langages qu'on déterminera et en déduire que L est nécessairement rationnel.